White Paper

Report ID: 2879327

Application Number: HG-229349-15

Project Director: Kathryn M. Tomasek

Institution: Wheaton College

Reporting Period: 5/1/2015-9/30/2016

Report Due: 12/31/2016

Date Submitted: 2/24/2017

White Paper
Ref: HG-229349-15

MEDEA: Modeling semantically Enhanced Digital Edition of Accounts – Digital Scholarship for the Semantic Web
Kathryn Tomasek, Wheaton College, Massachusetts[1]

Introduction

A significant portion of the historical record, for which cost-effective digitization models have yet to be developed, includes accounts of various sorts that have been produced by most human communities. Books of accounts document numerous activities of everyday life: exchanges of goods and services; income and expenditures for households, businesses, institutions, and political entities; and payments for taxes, rental of property, use of transport, or rights-of-way. And extracting information from accounts has contributed to the development of long and well-respected traditions of economic and social history in both Europe and the United States.

MEDEA (Modeling semantically Enhanced Digital Edition of Accounts) lead researchers ask a new question: What if we treat accounts like humanities sources? That is, what if—instead of using social science models to extract information from accounts in order to apply statistical and computational methods for their analysis—we instead produce digital scholarly editions as a first step? That is, what if we transcribe books of account and mark them up as digital scholarly editions to be published on the web? Social science researchers could still extract information for statistical analysis, and at the same time we would have a growing set of data that could be used in ways that may not yet be developed, ways that might take advantage of the graph-based technologies of the semantic web. MEDEA lead researchers have posited that a critical mass of digital scholarly editions of historical accounting documents could open up historical scholarship to new levels of inquiry, and co-PI GeorgVogeler has been developing a bookkeeping ontology that can be used to expose individual sets of such editions to research using the affordances of the semantic web.

Since creating a critical mass of scholarly editions requires a critical mass of scholars to produce them, one significant focus of the activities during our award has been to expand the community of practice for digital scholarly edition of accounts. To this end, two workshops were held at the two host institutions: at the University of Regensburg in October 2015 and at Wheaton College in Massachusetts in April 2016. In this context we have extended the work on data modeling that Tomasek and Vogeler had begun independently in North America and in Europe, building a community of historical researchers in Europe, Asia, and North America who will contribute both to Vogeler's developing ontology for semantic markup of scholarly editions of accounts and to a growing collection of such editions that is housed at the Geisteswissenschaftliches Asset Management System (GAMS) at the Centre for Information Modeling – Austrian Centre for Digital Humanities at the Karl Franzens University in Graz.

Researchers at the three cooperating institutions—the University of Regensburg, the University of Graz, and Wheaton College in Massachusetts—have focused on an array of outcomes from the first year of the MEDEA cooperation. Professor Mark Spoerer and his student Kathrin Pindl have leveraged the exceptionally extensive accounts of the St. Catherine's Hospital in Regensburg, which was established in the twelfth century, to produce new price series that will be published under the auspices of the Chair of Economic and Social History at the University of Regensburg. Vogeler has demonstrated the utility of

[1] I am grateful to co-PIs Kathrin Pindl, a doctoral student under the Chair of Social and Economic History at the University of Regensburg, and Georg Vogeler, University Professor at the Centre for Information Modeling—Austrian Centre for Digital Humanities at the University of Graz, for their cooperation in leading the activities undertaken during the period of this award. Sections of this whitepaper entitled "Environmental Scan" and "Methodology and Standards" are drawn from the grant proposal that the three of us wrote in collaboration in the fall of 2014.

semantic web technologies—RDFs/OWL, SKOS, and SPARQL—for comparing data from several projects that have presented economic content from the medieval and early modern periods on the web (Vogeler 2016). My own contribution—in addition to extending the MEDEA community of practice beyond Europe and the United States—has been to collaborate with Wheaton College students in completing a full transcription of a local account book and to apply Vogeler's bookkeeping ontology for publication in the GAMS repository.

Thus, the main focus of this paper is a discussion of Vogeler's ontology, its relationship to the "transactionography" that my co-author Syd Bauman and I developed as a result of an earlier award (Tomasek and Bauman, 2013), and to present the application of Vogeler's ontology to the Daybook of Laban Morey Wheaton, one of a set of account books kept by a businessman in Norton, Massachusetts, between 1828 and 1859.[2] Examples of transcription and markup of the daybook are presented for the benefit of scholars interested in producing their own digital scholarly editions of historical accounting documents.

Environmental Scan

The current state of digital publications of financial records demonstrates the need for an integrated approach, and a significant goal of this stage of the bilateral MEDEA cooperation has been to bring together scholars working on the most significant projects to critique preliminary models proposed by Vogeler and Tomasek and to do additional data modeling. Certainly social science historians have been producing data sets for at least the past half-century, and these are stored in such repositories as the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan in Ann Arbor <https://www.icpsr.umich.edu>. Efforts at open access and conversion for interoperability are underway through the DDI Alliance <http://www.ddialliance.org/> and OpenICPSR <https://www.openicpsr.org/>. And American Historical Association Past President Patrick Manning is the lead researcher on the Collaborative for Historical Information and Analysis (CHIA) <http://www.chia.pitt.edu/>, a project that aims to compile historical data into a Human System Data Resource (Manning 2017).

The MEDEA project approaches the creation of data from a different direction, exploring opportunities for the production of richer and potentially more accurate data with recommendations that emerge from the practices of scholarly editing within the humanities. Ideally, MEDEA examples will integrate visual representation with both transcriptions marked with semantic values and tools for numerical analysis, as was accomplished with the accounts of the city of Basel in a project led by Susannah Burghartz (University of Basel) in collaboration with Georg Vogeler (University of Graz).[3]

Numerous European projects have produced and displayed data extracted from historical accounting documents, but to date only one project has presented a scholarly edition that fully integrated transcriptions with images of the original documents, the Alcalà account book project. This cooperation between the National University of Ireland at Maynooth and the University of Alcalà created a rich interface, including translations of the text with the help of proprietary technologies (Adobe Flash), but the project did not focus on data modeling. The full project is no longer online, and only the HTML pages are accessible via the Wayback Machine at archive.org.

Some visual representations of financial records have included far more than images of the pages of account books. The Hearth Tax Online <http://www.hearthtax.org.uk/index.html> displays maps that

[2] Encoding Historical Records for Historical Researach, Ref: HD-51224-11
[3] <http://gams.uni-graz.at/context:srbas>

express statistical analysis of data contained in major seventeenth-century English tax accounts as well as lists of surnames found in such regions as County Durham and the North, East, and West Ridings of Yorkshire, with search capabilities for London and Middlesex. The digital version of the Registers of the Comédie-Française <http://web.mit.edu/hyperstudio/cfr/> demonstrates the potential of database organization for producing visualizations based on statistical analysis. Lead investigators from these projects, Andrew Wareham and Kurt Fendt participated in the activities of the bilateral MEDEA project.

Other projects in Europe and the United States have focused on the textual representation of accounts. Such online critical editions include numerous examples of medieval accounts. Some of these include a 2007 digital version of a 2006 print critical edition: *Rotulus de denariis* by Thomas Jarry <http://www.unicaen.fr/mrsh/craham/revue/tabularia/print.php?dossier=dossier5&contribDebat=true&file=03jarry_nouvelleed.xml>. R. Anthony Lodge took a similar approach in his publication of Les comptes des consuls de Montferrand (1273–1319) <http://elec.enc.sorbonne.fr/montferrand/>. And on a much larger scale the Comptes de châtellenies savoyarde <http://www.castellanie.net> project in Chambéry (Christian Guilleré) and Lyon (Jean-Louis Gaulin, Marjorie Burghardt) created a resource presenting images and transcriptions of account books from Northwestern Italy and Southeastern France. This latter edition employs an image viewer that requires Adobe Flash for access to page images.

The MEDEA project has drawn on German examples as well. Before 2009, Otto Volk's *Computatio* was the major German resource on medieval and early modern accounting documents with a huge bibliography and a glossary of special terms.[4] Researchers at the University of Mainz have recently begun work on a digital edition of the Augsburg Baumeisterbücher <http://www.historische.kulturwissenschaften.uni-mainz.de/1089.php>. Principal investigator Jörg Rogge participated in the workshops and projects in the first stage of the bilateral MEDEA project.

Small instances of account transcriptions can be found in some projects in the United States. Examples include the Alexander D. Betts' Account Book of Student Expenses, 1852-1853 <http://docsouth.unc.edu/unc/unc06-53/unc06-53.html> in the Documenting the American South project as well as images and transcriptions of a few extracts from business ledgers in the Bethlehem Digital History Project <http://bdhp.moravian.edu/community_records/business/busact.html>. Examples of fragmentary data from account books can also be found in the Railroads in the Making of Modern America project <http://railroads.unl.edu/topics/transcontinental.php>. Library professional Paige Morgan, who has undertaken the construction of a database of prices mentioned in English literature for her Visible Prices <http://www.paigemorgan.net/visibleprices/project-intro/> website, participated in MEDEA activities.

By far the most thorough efforts at modeling, transcribing, and representing digital editions of accounts books in the United States have been undertaken by documentary edition projects. One example is the Thomas Jefferson Farm Book and Memorandum Book, held and digitized by the Massachusetts Historical Society <http://www.masshist.org/thomasjeffersonpapers/>. Such projects tend to include accounts in their general methods of transcription-oriented publication. Over the past several years, editors of the Papers of George Washington <http://gwpapers.virginia.edu/> have begun to focus particularly on financial records (Stertzer 2014).

And finally, at least two European projects support the semantic analysis of the accounts edited: Die mittelalterlichen Schuld- und Rechnungsbücher des Deutschen Ordens um 1400 <http://www.schuredo.uni-hamburg.de/content/> edited at Hamburg University under the lead of Jürgen Sarnowsky, and the Henry III fine rolls <http://www.finerollshenry3.org.uk/home.html> published at the

[4] Recent efforts to access the site at <http://online-media.uni-marburg.de/ma_geschichte/computatio/> have resulted in a permissions error message.

Department of Digital Humanities at King's College London. Both have focused their semantic enrichment on lists of persons, places, and goods mentioned in the documents, with only a minor interest in the numerical interpretation of the data that can be extracted from the transcriptions. In contrast, the team of Burghartz and Vogeler attended closely to numbers in their compilation of the sixteenth-century account books of the city of Basel (http://gams.uni-graz.at/context:srbas). This edition uses a combination of TEI and Semantic Web technologies.

Methodology and Standards

Less well researched is the data modeling that would help to create scholarly, verifiable, and exchangeable data taken from account books. Contemporary spreadsheet software appears to offer a transparent representation of the same kinds of numerical information found in accounts recorded on paper spreadsheets and ledgers. Since these digital representations of accounts as relational databases emphasize the ability to perform calculations, they privilege numerical data and offer little utility with regard to representations of semantic information. The apparent uniformity maintained in this model of account information flattens out the data and loses much of richness of the original documents.

One main reference point for the development of standards for the production of semantically enriched data from analog originals is the Guidelines of the Text Encoding Initiative (TEI), an international standard for markup of the sorts of printed and manuscript texts that constitute the raw materials for literary, linguistic, and historical studies. Numerous digitization projects over the past twenty years have used TEI conformant transcription, but accounts have been neglected by most projects focused on creating scholarly editions until the past few years. The current version of the TEI Guidelines offers models for representing contextual or semantic information in files associated with transcriptions of analog sources; these include prosopographies, gazetteers, lists of organizations (sometimes informally called "orgographies"), and "transactionographies." As Tomasek and Bauman have demonstrated for nineteenth-century English-language account books and Vogeler has shown for medieval and early modern accounts from Europe, the established methods for text markup contained in the TEI Guidelines are insufficient for representing the wealth of information to be found in accounts of various genres.

In both cases, however, the recommendation has been to begin with transcription and markup of the sources using the TEI Guidelines since they are a longstanding international standard for preparing digital scholarly editions of historical texts and manuscripts for interchange and for multiple uses. The "transactionography" model presented by Tomasek and Bauman is a TEI customization that represents the core element of historical account books—the transaction—as a set of "transfers" between people or organizations. Since "from" and "to" have assigned meanings associated with the `<date>` element in TEI, the directions of exchange of currency, goods, services, rights, or credit are indicated using "fra" and "til"—the translations for "from" and "to" in Norwegian, as suggested by a participant in an earlier set of meetings under the previous award referenced above.

Vogeler has recommended steps beyond transcription and markup that look to another reference point as we explore data models for digital editions of accounts, suggesting a turn to opportunities presented by the Semantic Web, which contribute to growing interest in the Digital Humanities community in building ontologies for exchanging data. This is based on the W3C proposal for creating exchangeable data resources with RDF <http://www.w3.org/RDF/> and modeling them with the Web Ontology Language OWL <http://www.w3.org/standards/techs/owl#w3c_all>. One of the major humanities initiatives in this field is the CIDOC-CRM <http://www.cidoc-crm.org/>. Using this conceptual reference model, the museum community seeks to develop an upper level ontology for cultural heritage objects. One of the authors of the CIDOC_CRM, Øyvind Eide, participated in MEDEA activities, working with Vogeler on the MEDEA ontology and mapping its relationship to the content reference model. A preliminary version

of the MEDEA bookkeeping ontology is available on github: <https://github.com/GVogeler/bookkeeping>.

"Transactionography" and the MEDEA Bookkeeping Ontology

"Transactionography," the TEI customization for account books developed by Tomasek and Bauman employs the standoff markup available in TEI P5 to propose what is fundamentally an XML serialization of information about the human activities that are recorded in historical accounting documents. Such an XML serialization is not unlike the RDF serialization that Vogeler recommends as a step for exposing the information in account books for research on the semantic web. In fact, Vogeler and Eide concluded in as yet unpublished research that the "transactionography" customization parallels portions of the MEDEA ontology and that both are compatible with the CIDOC-CRM data model. Futhermore in an article based on research performed during the period of the MEDEA award, Vogeler has demonstrated that the MEDEA bookkeeping ontology can facilitate comparison of bookkeeping information from multiple projects displaying information from historical account books on the web (Vogeler 2016).

Vogeler teaches workshops in which he recommends use of the `@ana` attribute to add references to the MEDEA bookkeeping ontology to the TEI markup of digital scholarly editions of accounts. The addition of such references enables the extraction of an RDF serialization that then allows for comparison of the data on the semantic web. The references used with the `@ana` attribute include the following:

```
#bk_entry
#bk_amount
#bk_quantity
#bk_what
#bk_to
#bk_from
#bk_debit
#bk_credit
#bk_between
#bk_transfer
```

The sample markup in the following section demonstrates use of these references to the MEDEA ontology in transcriptions from a nineteenth-century daybook.

Sample Markup from the Wheaton Family Papers

In early summer 2016, four Wheaton College students spent eight weeks completing a full transcription of the daybook kept by local businessman Laban Morey Wheaton between 1828 and 1859. Wheaton, a member of the family that founded Wheaton Female Seminary (which later became Wheaton College), kept a general store, rented lands, and owned textile mills in Norton, Massachusetts, during the second quarter of the nineteenth century. Accounting records associated with some of these businesses are held in the Wheaton College Archives and Special Collections, and they have been the focus of the Wheaton College Digital History Project since 2009. These records document daily economic activities in a New England town undergoing a transition from an agricultural economy to an industrial one. Examples from the first two pages of the daybook demonstrate the simplicity of including markup references to the MEDEA bookkeeping ontology in preparation for publication of this edition in the GAMS repository at the University of Graz.

The daybook's flyleaf offers the first piece of information to be marked with a reference to the ontology. Since a daybook lists *transactions* involving *transfers* between a businessman and his clients, the markup for the information in the first line of the flyleaf (Fig. 1) is

```
    <p>
    <name type="person" ref="#pers_WCDH002" ana="#bk_from">Laban M.
Wheaton's Day Book</name>
    <choice>
     <abbr>C.</abbr>
     <expan>Creditor</expan>
    </choice></p>
```

The <name> element includes attributes that describe the @type of name, the @ref reference to an XML:id found in a prosopography of persons named in the documents in the archival collection of papers created by members of the Wheaton family, and an @ana reference to the MEDEA bookkeeping ontology. In the latter, #bk_from describes Wheaton as one of the people involved in the *transfers* that compose the *transactions* recorded in the daybook.

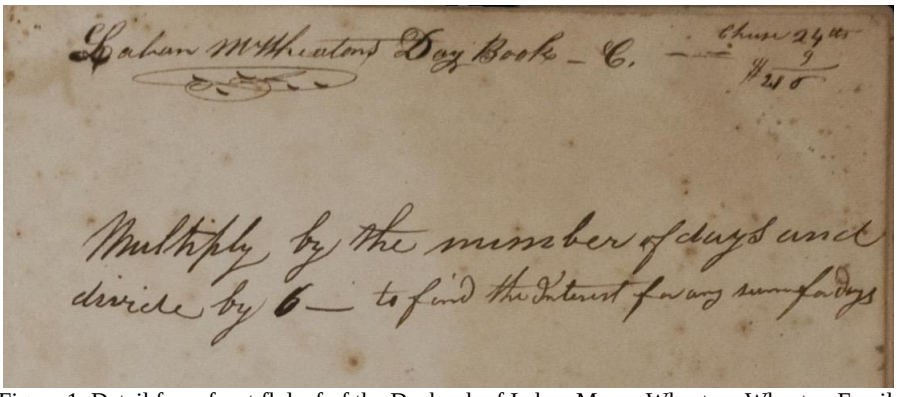

Figure 1. Detail from front flyleaf of the Daybook of Laban Morey Wheaton. Wheaton Family Papers. MC001. Wheaton College Archives and Special Collections. Wheaton College. Norton, Massachusetts.

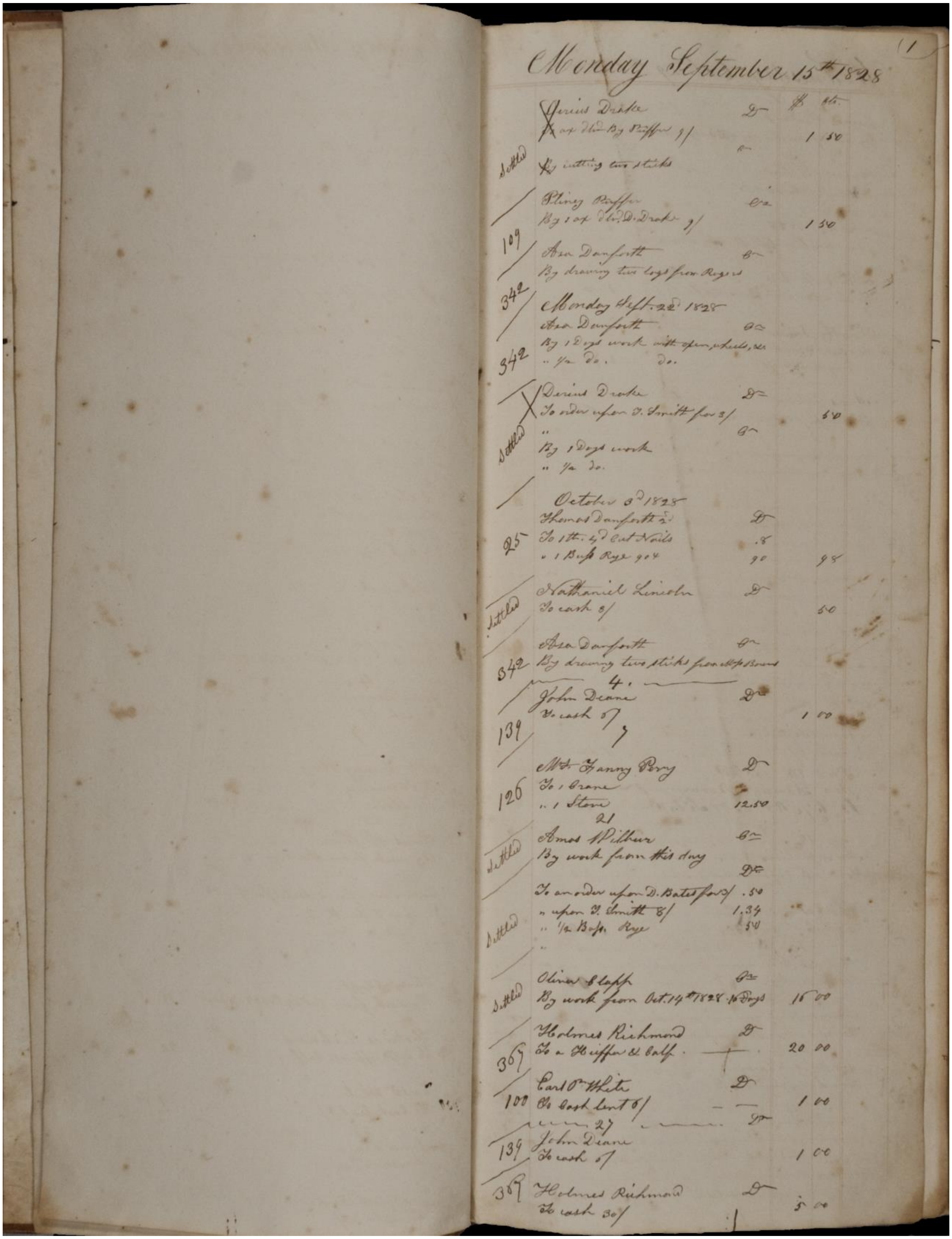

Figure 2. First page of the Daybook of Laban Morey Wheaton. Wheaton Family Papers. MC001. Wheaton College Archives and Special Collections. Wheaton College. Norton, Massachusetts.

Additional references to the MEDEA bookkeeping ontology are illustrated in markup embedded in the transcription of the first page of the daybook (Fig. 2).

```
<pb n="001"
facs="https://digitalrepository.wheatoncollege.edu/bitstream/handle/11
040/17982/lmwdbk_000.001.jpg"/>
   <fw type="pageNum">001</fw>
   <table xml:id="p001" cols="6">
    <head>
     <span ana="#bk_date" from="#p001_01" to="#p001_03"> Monday<date
when="1828-09-15">September-15th-1828</date>
     </span>
     <table>
      <row>
       <cell/>
       <cell/>
       <cell>$</cell>
       <cell>cts.</cell>
       <cell/>
       <cell/>
      </row>
     </table>
    </head>
```

The Wheaton College Digital History project uses the <table> element to represent the layout of the pages. The XML:ids assigned to tables and rows are for use in extracting data for purposes that are not under discussion in this paper (constructing a transactionography). The example above demonstrates use of the element to indicate that dates noted in the daybook often refer to more than one transaction. Here, the element includes the @ana attribute to mark a "#bk_date"reference to the MEDEA bookkeeping ontology code for indicating *when* the *transactions* took place.

```
<row xml:id="p001_01" ana="#bk_entry">
     <cell>Settled</cell>
     <cell>
      <p><name type="person" ref="#pers_WCDH276" ana="#bk_to
#bk_debit">Derius Drake </name>
       <choice rend="align(right)">
        <abbr>D</abbr>
        <expan>debtor</expan>
       </choice></p>
      <p> To <measure commodity="ax" quantity="1">ax</measure>
       <choice>
        <abbr>dlvd</abbr>
        <expan>delivered</expan>
       </choice> By <name type="person" ref="#pers_WCDH524"/>Puffer 9/
      </p>
      <p>
       <name type="person" ref="#pers_WCDH276" ana="#bk_to #bk_credit
      #bk_what"/> By cutting two sticks
       <choice rend="align (right)">
```

```
              <abbr>Cr</abbr>
              <expan>Credit</expan>
             </choice>
            </p>
           </cell>
           <cell>
            <measure commodity="currency" quantity="1" unit="dollars"
      ana="#bk_amount">1</measure>
           </cell>
           <cell>
            <measure commodity="currency" quantity="50" unit="cents"
      ana="#bk_amount">50</measure>
           </cell>
           <cell/>
           <cell/>
          </row>
```

The first two transactions in the daybook indicate a series of exchanges among Laban Morey Wheaton and two men who worked for him. Wheaton did not extend credit to the first of these men listed, Derius Drake. In place of a reference to a page in a ledger where Wheaton would have recorded credit extended to Drake, we find the word "Settled." The entry indicates a full *transaction* involving two *transfers* in the form of debit and credit (i.e. commodity exchanged for work of equal value). Thus, the `<row>` element includes the `#bk_entry` and the `#bk_transfer` and the first use of the `<name>` element both the `#bk_to` and the `#bk_debit` references to the MEDEA ontology. The debit is balanced by the credit that Drake received for the work of cutting two sticks; thus the account is settled. The second use of the `<name>` element takes the `@ana` attributes `#bk_to` and `#bk_credit` as well as `#bk_what` to refer to the work of cutting the sticks. The final version of this markup will include a reference to a taxonomy of kinds of work that are mentioned in the daybook.

```
    <row xml:id="p001_02" ana="#bk_entry #bk_transfer">
     <cell ref="#ledger_p109">109</cell>
     <cell>
      <p><name type="person" ref="#pers_WCDH524" ana="#bk_to
#bk_credit #bk_what">Pliny Puffer</name>
       <choice rend="align (right)">
        <abbr>Cr</abbr>
        <expan>Credit</expan>
       </choice>
      </p>
      <p> By 1 ax <choice>
        <abbr>dlvd.</abbr>
        <expan>delivered to</expan>
       </choice>
       <name type="person" ref="#pers_WCDH276">Drake</name> 9/ </p>
     </cell>
     <cell>
      <measure commodity="currency" quantity="1" unit="dollars"
ana="bk_amount">1</measure>
     </cell>
     <cell>
      <measure commodity="currency" quantity="50" unit="cents"
```

```
ana="#bk_amount">50</measure>
     </cell>
     <cell/>
     <cell/>
    </row>
    <row xml:id="p001_03" ana="#bk_entry #bk_transfer">
     <cell ref="#ledger_p342">342 </cell>
     <cell>
      <p><name type="person" ref="#pers_WCDH255" ana="#bk_to
#bk_credit #bk_what">Asa Danforth</name>
       <choice rend="align (right)">
        <abbr>Cr</abbr>
        <expan>Credit</expan>
       </choice>
      </p>
      <p> By drawing two logs from <name type="person"
ref="#pers_WCDH540">Rogers</name>
      </p>
     </cell>
     <cell/>
     <cell/>
     <cell/>
     <cell/>
    </row>
```

The two following transactions on the page continue the use of references to the MEDEA ontology as established. Since Wheaton did extend credit to both Pliny Puffer and Asa Danforth, the first `<cell>` element for each transaction includes references to pages in the ledger in which Wheaton tracked the credit accounts of many of the people with whom he did business. Since the Puffer transaction documents his work of delivering an ax to Drake and the ax has already been referenced in the previous transaction, the `<name>` element that refers to Puffer includes in the `@ana` attribute the `#bk_what` reference to the MEDEA ontology and an `@ref` attribute will refer to a taxonomy of kinds of work. The same goes for Danforth's work of drawing two logs from Rogers.

Subsequent entries in the daybook indicate that Wheaton noted work sometimes as individual tasks like drawing logs or sticks, sometimes as a commodity for which he paid people by the day, and sometimes as orders drawn on other people to whom he extended credit. In these cases, the markup differs accordingly.

```
<row xml:id="p001_04" ana="#bk_entry #bk_transfer">
     <cell>342</cell>
     <cell><p><date ana="#bk_date" when="1828-09-22">Monday Sept. 22d
1828 <span from="#p001_04" to="#p001_06"/>
       </date></p>
      <p><name type="person" ref="#pers_WCDH255" ana="#bk_to
#bk_credit #bk_what">Asa Danforth</name>
       <choice rend="align (right)">
        <abbr>Cr</abbr>
        <expan>Credit</expan>
       </choice>
      </p>
      <p> By <measure quantity="1" unit="days" commodity="work">1 Days
```

```
work with oxen, wheels,
        &amp;c</measure>

      </p>
      <p>
       <measure quantity=".5" unit="days" commodity="work">" 1/2 do.
do.</measure>
      </p>
     </cell>
     <cell/>
     <cell/>
     <cell/>
     <cell/>
    </row>
    <row xml:id="p001_05" ana="#bk_entry #bk_transfer">
     <cell>Settled</cell>
     <cell>
      <p><name type="person" ref="#pers_WCDH276" ana="#bk_to #bk_debit
#bk_what">Derius Drake</name>
       <choice rend="align(right)">
        <abbr>D</abbr>
        <expan>debtor</expan>
       </choice>
      </p>
      <p> To order upon
     <name type="person" ref="#pers_WCDH574">T. Smith</name> for 3/
     </p>
          </cell>
          <cell/>
          <cell>
           <measure commodity="currency" quantity="50" unit="cents"
     ana="#bk_amount">50</measure></cell>
          <cell/>
          <cell/>
         </row>
         <row xml:id="p001_06" ana="#bk_entry #bk_transfer">
          <cell>Settled</cell>
          <cell>
           <p>
            <name type="person" ref="#pers_WCDH276" ana="#bk_to
     #bk_credit #bk_what">"</name>
            <choice rend="align (right)">
             <abbr>Cr</abbr>
             <expan>Credit</expan>
            </choice>
           </p>
           <p> By<measure quantity="1" unit="Days" commodity="work">1
     Days work</measure>
           </p>
           <p> " <measure quantity=".5" unit="days" commodity="work">"
     1/2 do.</measure>
           </p>
```

```
          </cell>
          <cell/>
          <cell/>
          <cell/>
          <cell/>
         </row>
```

In the first transaction above the `<name>` element includes the `#bk_what` reference to the MEDEA ontology and will also include an `@ref` attribute to a taxonomy of kinds of work, in this case, a teamster's work with oxen. Once again, Wheaton's notation indicates that he carries an account for Danforth but not for Drake, whose day and a half of work is settled through an account that Wheaton does carry for Smith.

The rest of the transactions listed on the page include additional references to work as well as exchanges of cash or credit for commodities from four-penny nails and a bushel of rye to a heifer and calf to a crane and stove. All of these commodities will need to be listed in a taxonomy to be referenced in additional markup.

Numerous references to Wheaton's cash account as well as occasional mentions of money lent to individuals indicate the need for additional taxonomies. The first example below demonstrates markup for the cash account; the second an instance of money lent. For the loan, `#bk_what` refers to the MEDEA bookkeeping ontology, and a pointer to a taxonomy will be needed.

```
    <row xml:id="p001_08" ana="#bk_entry #bk_transfer">
     <cell>Settled </cell>
     <cell>
      <p><name type="person" ref="#pers_WCDH713" ana="#bk_to
#bk_debit">Nathaniel Lincoln</name>
      </p>
      <p> To <name type="person" ref="#pers_WCDH002"
ana="#bk_credit">cash</name> 8/ </p>
     </cell>
     <cell/>
     <cell>
      <measure commodity="currency" quantity="50" unit="cents"
ana="#bk_amount">50</measure></cell>
     <cell> </cell>
     <cell> </cell>
    </row>

    <row xml:id="p001_16" ana="#bk_entry #bk_transfer">
     <cell>100 </cell>
     <cell>
      <p>
       <name type="person" ref="#pers_WCDH656" ana="#bk_to #bk_debit
#bk_what">Earl O White</name>
       <choice rend="align(right)">
        <abbr>D</abbr>
        <expan>debtor</expan>
       </choice>
```

```
      </p>
      <p> To <name type="person" ref="#pers_WCDH002"
ana="#bk_credit">Cash</name> lent 6/ </p>
     </cell>
     <cell>
      <measure commodity="currency" quantity="1" unit="dollars"
ana="#bk_amount">1</measure></cell>
     <cell>
      <measure commodity="currency" quantity="00" unit="cents"
ana="#bk_amount">00</measure></cell>
     <cell/>
     <cell/>
    </row>
```

These examples of markup that include references to the MEDEA ontology indicate that the ontology itself remains insufficient to express all of the information a researcher might wish to mark in a digital scholarly edition of an account book. Taxonomies for kinds of work and for commodities as well as for loans or for cash debited to the businessman's cash account will be needed as well. Vogeler has demonstrated that it is possible to build on the work of previous scholars to construct taxonomies of goods that can be applied across regions and languages in medieval and early modern examples. Using the MEDEA bookkeeping ontology in conjunction with such taxonomies can then be used to produce RDF triple stores that can expose the data from historical accounts to discovery on the Semantic Web (Vogeler 2016).

Conclusion

Thus the data modeling, transcription and markup, and research accomplished during MEDEA's first year have advanced the work of extending the community of practice, developing the MEDEA ontology, and testing its application in several instances. In addition, Vogeler's experiences with the Basel accounts and Tomasek's with the Wheaton daybook offer evidence that the work is not overly labor intensive and can be completed within reasonable amounts of time. In the Basel example, two researchers with graduate level expertise completed the transcription and markup of hundreds of years' worth of accounts in a single year. In the Wheaton College example, four undergraduates completed the transcription and markup of a daybook of more than two hundred pages in eight weeks.

Creating additional digital scholarly editions of accounts with the relatively straightforward inclusion of references to the MEDEA ontology will facilitate further research into the possibilities for comparison of bookkeeping information that can then be extracted from the accounts. And such possibilities represent only one set of research questions that can be supported with digital scholarly editions of accounts. Additional possibilities include examination of networks of interaction and exchange at the local level as well as at regional and perhaps even global levels. At the granular local level an examination for the town of Norton, Massachusetts, might expand our historical understanding of communities undergoing the transition from agricultural to industrial economies. Web publication of the files associated with digital scholarly editions of the multiple account books in a single local collection could certainly be used to help undergraduates understand the incremental pace of such transitions and the impacts of changing technologies in the past.

Further development and testing will be required to explore more fully the research opportunities that might be opened up if more scholars create digital scholarly editions with an eye to interoperability. At this point, presentation of high-quality images of the documents alongside digital scholarly editions with

markup referencing the MEDEA bookkeeping ontology is an important first step that more historical researchers are beginning to undertake in Asia, Europe, and North America.

Bibliography of Works Cited

Guidelines for Electronic Text Encoding and Interchange, originally edited by C.M. Sperberg-McQueen, Lou Burnard, version 2.9.1, 2015 http://www.tei-c.org/Guidelines/P5/.

Kokaze, N.; Nagasaki, K.; Shimoda, M.; Muller, A.C. (2016). Modeling New TEI/XML Attributes for the Semantic Markup of Historical Transactions, based on 'Transactionography.' Proceedings of the 6th Conference of Japanese Association for Digital Humanities: "Digital Scholarship in History and the Humanities. Tokyo: University of Tokyo, 2016. http://conf2016.jadh.org/abstracts/s5-2/

Manning, Patrick. 2017. Inequality: Historical and Disciplinary Approaches. American Historical Review, 122 (1): 1-22. DOI: https://doi.org/10.1093/ahr/122.1.1 Published: 31 January 2017. Accessed: 17 February 2017.

Stertzer, Jennifer. 2014. Working with the Financial Records of George Washington. Document vs. Data, in: Digital Studies 4 <http://www.digitalstudies.org/ojs/index.php/digital_studies/article/view/247/308>.

Tomasek, Kathryn, and Syd Bauman. 2013 Encoding Financial Records for Historical Research, in: jTEI 6 <http://jtei.revues.org/895>, DOI : 10.4000/jtei.895. Japanese translation by Naoki Kokaze: <http://hdl.handle.net/2261/56940>.

Vogeler, Georg. 2016. The Content of Accounts and Registers in their Digital Edition: XML/TEI, Spreadsheets, and Semantic Web Technologies. In *Konzeptionelle Überlegungen zur Edition von Rechnungen und Amtsbüchern des späten Mittelalters*, ed. Jürgen Sarnowsky. V&R Unipress.